Title

Protocol-processing overhead on the performance of error recovery schemes in high-speed network environments

Authors

Bae, Jaime Jungok
Suda, Tatsuya
Watanabe, Naoya

Publication Date

1991

Peer reviewed

Protocol-Processing Overhead on the Performance of Error Recovery Schemes in High-Speed Network Environments

Jaime Jungok Bae, Tatsuya Suda and Naoya Watanabe

Technical Report No. 91-47

Department of Information and Computer Science
University of California, Irvine
Irvine, CA 92717

Protocol-Processing Overhead on the Performance of Error Recovery Schemes in High-Speed Network Environments*

Jaime Jungok Bae and Tatsuya Suda
Department of Information and Computer Science
University of California, Irvine
Irvine, CA 92717
U.S.A.
(phone) 714-856-5474
(fax) 714-856-4056

Naoya Watanabe
Research and Development Planning Group
NTT Research and Development Headquarters
NTT, Yamate Life Insurance Bldg., 10th floor
1-7, Uchisaiwai-Cho 1-Chome
Chiyoda-ku, Tokyo, 100, Japan
(phone) 011-81-3-509-2417
(fax) 011-81-3-595-4521

* This material is based upon work supported by the National Science Foundation under Grant No. NCR-8907909. This research is also in part supported by University of California MICRO program.

Abstract

This paper investigates the effects of protocol-processing overhead on the performance of error recovery schemes in high-speed network environments. The investigated error recovery schemes are:

- an edge-to-edge error recovery scheme, where retransmissions of erred packets only take place between source and destination nodes, and
- a link-by-link error recovery scheme, where retransmissions only take place between adjacent switching nodes.

For retransmission of erred packets, we consider both Go-Back-N and Selective-Repeat procedures in the analysis.

The performance measures we obtain are the distribution of transfer delays and the loss probability of packets across a network. To obtain these measures, this paper develops a tandem queueing network model with feedbacks where each queue represents a protocol layer within a switching node, rather than a switching node as a whole.

Numerical results show that for a network with very-high-speed/low-error-rate channels, an edge-to-edge scheme gives the smaller packet transmission delay than a link-by-link scheme for both Go-back-N and Selective-Repeat retransmission procedures, while keeping the packet loss probability sufficiently small.

1. Introduction

The arrival of the information age has intensified the demand for communication services of all kinds. Applications for voice and data, as well as video, are rapidly expanding. Hence, future communication services must be able to facilitate a wide variety of diversified services in a practical and easily expandable fashion [1]. In order to effectively handle a broad range of services, new network architectures, such as fast-packet-switching and Asynchronous Transfer Mode (ATM), have been proposed [2-4]. Such networks are characterized by very high transmission rates on network links and a simple, hardwired node-to-node protocol which matches the rapid channel speed of the network. The recent technological advances, especially in fiber optics and micro electronics, have made such changes possible.

The major advantage in adopting simplified protocol in a high-speed network is that it reduces the protocol-processing overhead at each switching node on a network, and thus reduces packet delays across the network. Protocol-processing time, such as time required to detect and possibly correct errors, is comparatively large in high-speed networks and becomes dominant in determining the packet transfer delay across a network. For instance, transmission time of a 1000 bit packet on a 1 Mbits/sec channel is 1 msec, whereas that of the same packet on a 1 Gbits/sec channel is 1 μsec, which makes protocol-processing time at a switching node comparatively large. Because it ensures a high quality data transport through a network at the expense of increased packet transfer delays due to processing overhead at each switching node, the current network architecture (e.g., X.25, ISO 7 layer architecture), which employs strict error control between adjacent switching nodes, may not be suitable for a very high-speed network.

In this paper, we investigate the effects of protocol-processing overhead on the performance of error recovery schemes. We focus on the error recovery scheme used in high-speed networks; namely, the scheme where retransmissions of erred packets only take place between source and destination nodes (edge-to-edge error recovery scheme). We obtain the Laplace transform for the distribution of the end-to-end packet transfer delay, considering processing time required for error recovery. We also evaluate the performance of an alternative error recovery scheme; namely, a scheme where retransmissions take place between adjacent nodes (link-by-link error recovery scheme) and compare the performance of an edge-to-edge scheme with this alternative.

The rest of the paper is organized as follows. In section 2, we describe the network architecture to be investigated. In section 3, we develop a queueing network model to evaluate error recovery schemes. In section 4, an approximate analysis of the model developed in section 3 is presented, and the end-to-end packet transfer delay is obtained. In section 5, numerical results are presented to show the performance trade-offs between the error recovery schemes. Finally, in section 6, concluding remarks are made.

2. Network Architecture To Be Investigated

To investigate the effects of processing time on the network performance, we assume the following hypothetical layering architecture for high-speed networks [5], and compare it with conventional architecture. We focus on only the lower three layers of the OSI reference model.

In our hypothetical layering architecture, the basic transport mechanism is provided by the lower three levels of the protocol (1-, 2-, and 3-lower) applied to transmissions across each network link. The packet transport sublayer (level 3-upper) provides edge-to-edge communication within the network between source and destination nodes. The higher level end-to-end functions discussed in OSI model (i.e., transport layer and above) appear as higher layers above this basic transport mechanism and are not discussed here. The function of each layer in our layering architecture is described below.

- Level 1 - Physical Layer (P-layer) specifies the electrical characteristics and representation of transmitted bits.
- Level 2 - Link Layer (L-layer) performs several functions necessary for successful transmission between network nodes. These functions include frame delimiting and bit pattern transparency. As a major departure from conventional architecture (e.g., HDLC), error recovery procedures are not included in this level.
- Level 3 - This layer consists of the following two sublayers.
  - Level 3 Lower - Packet Network Sublayer (PN-layer) is the lower sub-layer of the layer 3. Primary function provided here is the routing of packets.
  - Level 3 Upper - Packet Transport Sublayer (PT-layer) is the upper sub-layer of the layer 3 and performs edge-to-edge error recovery.

In order to reduce the protocol-processing overhead, error recovery between adjacent switching nodes is not performed in the above layering architecture. Instead, reliable data communication through the network is provided at the edge of the network (i.e., at layer 3 upper); any detected errors are corrected with a peer communication between the source and destination nodes.

Propagation delay, as well as time required for protocol-processing, is an important factor which decides network performance. As an example, consider two adjacent switching nodes, A and B, linked by a 100 Km cable. Assume the typical propagation delay time of 5 μsec per 1 Km of a cable. Node A starts transmitting a packet. It takes 500 μsec for the electric signal to propagate to node B. This propagation delay is significantly greater than the packet transmission time of 1 μsec in the previous example. Therefore, propagation delay, which is assumed to be negligible in existing networks, becomes another dominant factor in determining the packet transfer delay across a network. In our analysis, both processing time and the propagation delay are considered in obtaining the end-to-end packet transfer delay.

3. Analytic Model

3.1 Protocol Layer Model

This paper seeks to evaluate protocol-processing overhead in high-speed network environments, and thus each layer at a switching node, rather than the switching node as a whole, is modeled as a queueing system in our analytic model [6]. In modeling edge-to-edge and link-by-link error recovery schemes, rather than considering all the functions performed at each protocol layer, we focus on the packet transmission and error recovery functions. The protocol architectures we investigate are described below (see Table 1). In an edge-to-edge scheme, L-layer performs frame transmission, but hop-by-hop error recovery is not performed. PN-layer performs routing function. PT-layer performs end-to-end error recovery. In a link-by-link scheme, L-layer transmits frames and correct errors on a hop-by-hop basis. PN-layer routes packets, and PT-layer performs no specific function.

We assume that data units at L-layer and at PN- and PT-layers are of the same length, and thus in our analysis, we do not distinguish frames from packets. It is also worth noting that in the analysis, variables with subscript 1, 2, and 3 are associated with L-layer, PN-layer, and PT-layer respectively.

3.2 Queueing Network Model

Figure 1 shows the queueing network model for link-by-link and edge-to-edge error recovery schemes based on the protocol architecture described in the previous section. In both link-by-link and edge-to-edge schemes, higher layers pass new packets to PT-layer at a source station. We assume that new packets arrive at PT-layer according to a Poisson process with rate λ_3. We further assume that the packet length is exponentially distributed with the average P bits and hence, the average transmission time of a packet is P/V sec, where V is the speed of the physical channel.

Packets are stored and processed at PT-layer and then passed to PN-layer. Processing time at PT-layer of a transmitting node is assumed to be exponentially distributed with the rate $\mu_{3,t}$ (subscript t stands for "transmitting node processing time"). PN-layer makes routing decisions and passes packets to L-layer. Processing time at PN-layer is exponentially distributed with the rate μ_2. Note that since the same rate μ_2 is assumed at PN-layer throughout the network, subscript t is not necessary. At L-layer packets are stored and then transmitted through a physical transmission link. The time spent at a transmitting node L-layer consists of a packet transmission time, and thus it is exponentially distributed with the average $1/\mu_{1,t} = P/V$.

In a link-by-link scheme, L-layer stores incoming packets at an intermediate switching node and examines the packets for errors. Processing time to detect errors at L-layer is assumed to be exponentially distributed with the rate $\mu_{1,e}$ (subscript e stands for "error detection time"). If no error is detected, the receiver immediately sends an ACK back to the sender. This is indicated by a feedback line (an arrow) between two adjacent L-layers in Figure 1. In case of error, no ACK is sent to the sender, and the sender retransmits the packet after the specified link-by-link time-out period. The procedure for retransmitting erred packets is described in the next subsection. If no error is found at L-layer, the packet is passed to PN-layer, where a routing decision is made. If the packet is addressed to some other node, it is passed down to L-layer for the further transmission to the next node on the path to the destination. If the packet is destined for that node, it is passed up to PT-layer, and PT-layer immediately forwards the packet to the higher layer. Note that in a link-by-link scheme, no processing is done at PT-layer at destination.

In an edge-to-edge scheme, when an intermediate L-layer queue receives a packet, it immediately passes the packet to PN-layer for routing. L-layer does not perform any error checking on incoming packets (i.e., $\mu_{1,e} = \infty$). As in a link-by-link scheme, if the packet is addressed to some other node, PN-layer passes it down to L-layer for the further transmission to the next node. If the packet is destined for that node, PN-layer passes it up to PT-layer. In an edge-to-edge scheme, PT-layer at the destination node performs error checking, and the processing time to detect errors at PT-layer is assumed to be exponentially distributed with the rate $\mu_{3,e}$. If no error is detected, the destination node immediately sends an ACK back to the source node. This is indicated by a feedback line from a destination to a source node in the figure. If there is an error in a packet, the source node retransmits the packet after the specified edge-to-edge time-out period.

In the model, propagation delay along the link is also considered. For simplicity, constant propagation delay D_{prop} is assumed between adjacent nodes (i.e., internodal distance is constant). Note that there may be some interfering traffic at each switching node. The effect of this interfering traffic can easily be modeled by reducing the communication capacity (i.e., a service rate of each queueing system in our model) as in [7, 8].

3.3 Errors, Retransmissions, and Time-outs

In a link-by-link scheme, in case of error, a packet may be transmitted up to M_1 times (initial transmission and up to $M_1 - 1$ retransmissions) at L-layer between two adjacent nodes. In an edge-to-edge scheme, a packet may be transmitted up to M_3 times (initial transmission and up to $M_3 - 1$ retransmissions) at PT-layer between a source and a destination node. Note that in our analysis, M_1 and M_3 could be either finite or infinite. If M_1 is finite, a packet received at L-layer still has an error (or errors) with probability $p_1^{M_1}$ at the end of M_1 transmissions, where p_1 is the probability that a packet suffers an error (or errors) on a link. We assume that packets which still have errors after M_1 transmissions are discarded and will not be forwarded to the next node. Therefore, in the link-by-link scheme, packet loss probability ε across a network becomes

$$\varepsilon = 1 - (1 - p_1^{M_1})^l \tag{1}$$

where l is the number of hops between source and destination nodes. (If M_1 is infinite, this loss probability approaches zero.)

Similar to the link-by-link case, if M_3 is finite in an edge-to-edge scheme, a packet received at PT-layer of the destination node still has an error (or errors) with probability $P_3^{M_3}$ at the end of M_3 transmissions, where p_3 is the probability that a packet arrives at PT-layer of the destination node with an error (or errors). p_3 is given by $1 - (1 - p_1)^l$, where $(1 - p_1)^l$ is the probability that no error occurs in a packet in l number of hops. Assuming the packets which still have errors after M_3 transmissions are discarded at destination PT-layer, packet loss probability ε across a network in the edge-to-edge scheme becomes

$$\varepsilon = p_3^{M_3} = (1 - (1 - p_1)^l)^{M_3}. \tag{2}$$

(If M_3 is infinite, this error probability approaches zero.)

For retransmission of erred packets at L-layer (in a link-by-link scheme) and at PT-layer (in an edge-to-edge scheme), we consider both Go-Back-N and Selective-Repeat procedures in the analysis.

In Go-Back-N, a sender can send packets up to a given window size (N) without receiving acknowledgements from a receiver. When the receiver receives a packet with no errors, it immediately sends an ACK back to the sender. If an ACK is not received within a specified time-out period, the packet is assumed lost or erred, and the sender retransmits all the packets starting with the lost/erred packet. In Selective-Repeat procedure, when the time-out period expires, the sender retransmits only the lost/erred packet.

Constant time-out periods τ_3 and τ_1 are used for retransmissions in the edge-to-edge and link-by-link schemes, respectively (see Figure 2). The edge-to-edge time-out period begins immediately after PT-layer passes a packet to PN-layer. The link-by-link time-out period begins immediately after transmission of a packet is completed at the sender L-layer. In both the edge-to-edge and link-by-link schemes, it is assumed that a time-out occurs only when a packet is, indeed, lost or erred. In reality, improperly short time-out periods and long queueing delays within a network can cause time-outs even when packet transmission is successful. For simplicity, this case is not considered in the analysis (The same assumption is made in [8 - 11].)

We further define r_3 as the average time interval from an arrival of a packet at the source PT-layer queue to the end of the edge-to-edge time-out period, and r_1 as the average time interval from an arrival of a packet at a sender L-layer queue to the end of the link-by-link time-out period (see Figure 2). From Figure 2(a), it is easy to see that r_3 consists of the average system time (i.e., the average queueing delay plus the service time) at the source PT-layer queue and the edge-to-edge time-out period τ_3. From Figure 2(b), r_1 consists of the average system time (i.e., the average queueing delay plus packet transmission time) at a sender L-layer queue and the link-by-link time-out period τ_1. Note that the average queueing delay at a sender L-layer queue differs at each node since the arrival rate to each sender L-layer queue changes due to the packet discarding. (Packets which still have errors after M_1 transmissions are discarded.) Therefore, each node has different values of r_1. We let $r_{1,i}$ denote r_1 at node i.

4. Analysis

4.1 Assumptions

In order to make the analysis tractable, we employ the following assumptions in our analysis:

- Each time a packet joins a queue in our queueing network model, its length is determined afresh from an identical exponential distribution (with the average P bits).
- In an edge-to-edge scheme, the aggregated arrivals of new and retransmitted packets at a PT-layer source queue are assumed to follow a Poisson process. In a link-by-link scheme,

it is assumed that at each sender L-layer queue, the aggregated arrivals of packets passed from PN-layer and retransmitted packets follow a Poisson process.

The first assumption is a well-known independence assumption [12] and used in a number of literatures. The following explains the validity of the second assumption. This paper focuses on high-speed networks which utilize very high-speed and high-quality channels (i.e., optical fibers). Optical fibers can provide very low error probability (i.e., 10^{-9} to 10^{-13} bit error rate), and thus packet retransmissions due to errors are very scarce in such networks. Furthermore, as we will see in the numerical example section (Figure 6), in case of errors, almost all the errors are corrected by at most two transmissions (initial transmission and one retransmission), and thus packet discarding will rarely happen. Occasional retransmissions and losses will not severely destroy the Poisson property of an input process.

We further verify the accuracy of the second assumption through simulations. In our simulation, 2×10^{-7} is used for the bit error rate on a link. Due to the extensive memory and simulation run-time requirements in obtaining statistically meaningful data for a network with an extremely low error rate, 2×10^{-7} bit error rate was chosen, as opposed to more realistic bit rates such as 10^{-9}. If the accuracy of the Poisson assumption is established for a network with 2×10^{-7} bit error rate on a link, this assumption certainly holds for a network with smaller error rates such as 10^{-9}.

Figure 3 represents the queueing model used in the simulations for both link-by-link and edge-to-edge schemes. This one-hop network is simulated allowing the maximum of 2 transmissions per packet (initial transmission and one retransmission). We assumed Selective-Repeat retransmission procedure in our simulations.

Using the average packet transmission time as a unit time (i.e., $1/\mu_{1t} = 1$), the following parameter values are used: $\mu_{1,e} = \mu_2 = \mu_{3,t} = \mu_{3,e} = \mu_{1,t}$, $D_{prop} = 30$, and $\tau_3 = \tau_1 = 84$. (The validity of these parameter values is discussed in section 5.) The Poisson assumption is tested at two points, Point 1 and Point 2 in Figure 3. At Point 1, the aggregated arrivals of new and retransmitted packets are tested. At Point 2, the departure process after packet discarding takes place is tested. Chi-Square Goodness-of-Fit Test [13] is performed at these two points. As Table 2 indicates, the test results show that at the 5% level of significance, both the arrival pattern (at point 1) and the departure pattern (at point 2) appear to be Poisson.

In the above simulations, Selective-Repeat procedure is assumed. Go-Back-N retransmission procedure requires much more memory and run-time in simulations than Selective-Repeat procedure, thus severely limiting the values of a link error rate which can be used in simulations to very unrealistic values. Because of this, we could not test the Poisson assumption for the Go-Back-N case. However, due to the intuitive reasons mentioned earlier, we believe that the Poisson assumption holds for networks with a very low link error probability (e.g. 10^{-9} bit error rate) when the Go-Back-N retransmission procedure is employed.

In the analysis, we also assume that retransmissions do not have priority over transmissions of new packets and that errors do not occur in ACK packets. Further, it is assumed that each node has infinite buffers. In the case where Go-Back-N is used, infinite window size is assumed.

With the assumptions made in this section, our queueing network model becomes of the Jackson's type [14], and each queue in the network behaves as if it were an independent M/M/1 system. In subsections 4.2 and 4.3, a link-by-link scheme and an edge-to-edge scheme are analyzed, respectively.

4.2 Link-by-Link Scheme

In this section, we focus on a link-by-link error recovery scheme and analyze its performance.

Effective Bit Rate

We first consider a queueing system at L-layer of a sending node (i.e., a queue with the service rate $\mu_{1,t}$ in Figure 4.) Note that when Go-Back-N retransmission procedure is employed, the sending node retransmits all the packets in its L-layer queue starting with the lost/erred packet when the time-out period τ_1 expires. With this in mind, if the initial transmission of a packet fails at node i, at the first retransmission (i.e., the second transmission) of the erred packet, the packets which arrived during $r_{1,i}$ plus the erred packet itself are retransmitted (see Figure 2(b)). Thus, $r_{1,i}\lambda_{1,i}+1$ number of packets are transmitted at the second transmission of the erred packet, where $\lambda_{1,i}$ (Figure 4) is the packet arrival rate at L-layer at node i. At the j-th transmission of an erred packet, the packets that arrived during $(j-1)r_{1,i}$ are transmitted along with the erred packet. Letting $N_{1,i}$ denote $r_{1,i}\lambda_{1,i}$, a total of $(j-1)N_{1,i}+1$ packets are transmitted at node i at the j-th transmission of an erred packet when Go-Back-N procedure is employed. On the other hand, when Selective-Repeat procedure is employed, only the erred packet is retransmitted, and thus it is easy to see that $N_{1,i}=0$ for all i's.

Let $P(k)$ be the probability that a packet requires k transmissions (an initial transmission and $(k-1)$ retransmissions) before it is either passed up to PN-layer or discarded at the receiving L-layer (due to the limitation on the maximum number of retransmissions allowed), where $1 \le k \le M_1$. It is easily shown that $P(k)$ is given by

$$P(k) = \begin{cases} q_1 p_1^{k-1} & (1 \le k < M_1) \\ p_1^{M_1-1} & (k = M_1) \end{cases} \tag{3}$$

where $q_1 = 1 - p_1$. If a packet requires k transmissions (including the initial transmission), total $\sum_{j=1}^{k}\{(j-1)N_{1,i}+1\}$ number of packets are transmitted until the packet is accepted at the receiving L-layer. Therefore, the average number $Y_{1,i}$ of the packets to be transmitted until a packet is finally accepted at receiving L-layer becomes

$$\begin{aligned} Y_{1,i} &= \sum_{k=1}^{M_1}[\sum_{j=1}^{k}\{(j-1)N_{1,i}+1\}]P(k) \\ &= N_{1,i}\frac{p_1\{1 - M_1 p_1^{M_1-1} + (M_1-1)p_1^{M_1}\}}{(1-p_1)^2} + \frac{1-p_1^{M_1}}{1-p_1} \end{aligned} \tag{4}$$

From this we obtain the effective packet arrival rate $\lambda'_{1,i}$ (i.e., the aggregated arrival rate of the new packets from PN-layer and the retransmitted packets) at the sender L-layer queue at node i (Figure 4) as

$$\lambda'_{1,i} = \lambda_{1,i} Y_{1,i} = \lambda_{1,i}[N_{1,i}\frac{p_1\{1 - M_1 p_1^{M_1-1} + (M_1 - 1)p_1^{M_1}\}}{(1-p_1)^2} + \frac{1-p_1^{M_1}}{1-p_1}]. \tag{5}$$

We next consider the source node (see Figure 1). At the source node, we assume that packets arrive at PT-layer with the rate λ_3. Since the output from the PT-layer queue is the input to its PN-layer queue, the packet arrival rate λ_2 at the source PN-layer queue is $\lambda_2 = \lambda_3$. (Note that a link-by-link scheme is assumed in this section.) Also, since the output from the PN-layer queue is the input to its source L-layer queue, the packet arrival rate λ_1 at the sender L-layer queue from the PN-layer queue becomes $\lambda_1 = \lambda_2 (= \lambda_3)$. The packets from PN-layer (at the rate of λ_1) and retransmissions of erred packets (indicated by the feedback line in Figure 1) collectively form the arrival to the L-layer queue at the source node. The rate of this aggregated packet arrivals, λ'_1, is given by Eq.(5). Note that $\lambda_1 = \lambda_{1,1}$, $\lambda'_1 = \lambda'_{1,1}$, and $\lambda_2 = \lambda_{2,1}$.

At the intermediate nodes, since the packets with errors after M_1 transmissions are discarded at L-layer, the rate of packet arrivals at PN-layer from L-layer varies depending on the node (see Figure 4). Noting that $p_1^{M_1}$ is the probability that a packet is discarded (at node i) due to an error after M_1 transmissions, the packet arrival rate $\lambda_{2,i}$ at PN-layer at node i becomes

$$\lambda_{2,i} = \lambda_{1,(i-1)}(1 - p_1^{M_1}) \tag{6}$$

Further, the rate $\lambda_{1,i}$ at which packets are passed from PN-layer to sender L-layer at node i becomes $\lambda_{1,i} = \lambda_{2,i}$. Note that the rate of the aggregated arrivals of packets (packets from PN-layer at the rate of $\lambda_{1,i}$ and retransmissions due to errors) at sender L-layer is given by Eq.(5).

The effective utilization of each queue at node i becomes

- $\rho'_{1,i,t} = \lambda'_{1,i}/\mu_{1,t}$ at a sender L-layer queue at node i,
- $\rho'_{1,i,e} = \lambda'_{1,(i-1)}/\mu_{1,e}$ at a receiver L-layer queue at node i,
- $\rho_{2,i} = \lambda_{2,i}/\mu_2$ at a PN-layer queue at node i, and
- $\rho_{3,t} = \lambda_3/\mu_{3,t}$ at the PT-layer queue at the source node.

The utilization of the PT-layer queue $\rho_{3,e}$ at the destination is not considered since it is assumed that no processing is done at the destination PT-layer queue if link-by-link scheme is used.

End-to-End Packet Transfer Delay

We define the end-to-end transfer delay of a packet as the time starting when a packet first enters the source PT-layer queue and ending when that packet leaves the destination PT-layer. This transfer delay consists of (1) the time spent at a source PT-layer queue,

(2) the time spent at a source PN-layer queue, (3) the time spent at intermediate switching nodes, and (4) the time spent at the destination PN-layer queue. Note that the time spent at the destination PT-layer queue is not included since in a link-by-link scheme no processing is involved in PT-layer at the destination, and packets from PN-layer are immediately forwarded to the upper layer.

Let $B_3^*(s)$ be the Laplace transform for the distribution of the end-to-end packet transfer delay. $B_3^*(s)$ is obtained in the following way: let $F_{3,t}^*(s)$ be the Laplace transform for the distribution of the time spent by a packet in the source PT-layer queue, and $F_{2,i}^*(s)$ be the Laplace transform for the distribution of the time spent by a packet in a PN-layer queue at node i. Further, let $\Phi_{1,i}^*(s)$ be the Laplace transform for the distribution of the time required for a packet to "hop" to the next node, i.e., the time starting when a packet first enters the sender L-layer queue at node i and ending when the packet is passed to the PN-layer queue at node $i+1$, including the time incurred by retransmissions. Assuming that there are l hops from the source to the destination, $B_3^*(s)$ is given by

$$B_3^*(s) = F_{3,t}^*(s)F_{2,1}^*(s)\prod_{i=1}^{l}\{\Phi_{1,i}^*(s)F_{2,(i+1)}^*(s)\} \tag{7}$$

where $F_{3,t}^*(s)$ in the right hand side corresponds to the delay element (1) (see the definition of the transfer delay given in the above paragraph); $F_{2,1}^*(s)$ corresponds to (2); and the last product term $\prod_{i=1}^{l}\{\Phi_{1,i}^*(s)F_{2,(i+1)}^*(s)\}$ corresponds to (3) and (4).

We first obtain $\Phi_{1,i}^*(s)$ in Eq.(7). Let us consider L-layer queues in two adjacent nodes: a sender L-layer queue at node i (with the service rate $\mu_{1,t}$) and a receiver L-layer queue at node $i+1$ (with the service rate $\mu_{1,e}$) (see Figure 5). Let $F_{1,i,t}^*(s)$ and $F_{1,(i+1),e}^*(s)$ be the Laplace transform for the distribution of the time that a packet spends at the sender L-layer queue and at the receiver L-layer queue, respectively. As discussed in section 4.1, in a link-by-link scheme it is assumed that at sender L-layer, the aggregated arrivals of packets passed from PN-layer and retransmitted packets follow a Poisson process. (This implies that the input to the receiver L-layer queue at the next node also follows a Poisson process, since the service at the sender L-layer queue is exponential.) Therefore, we have

$$F_{1,i,t}^*(s) = \frac{\mu_{1,t}(1-\rho'_{1,i,t})}{s+\mu_{1,t}(1-\rho'_{1,i,t})} \tag{8}$$

$$F_{1,(i+1),e}^*(s) = \frac{\mu_{1,e}(1-\rho'_{1,(i+1),e})}{s+\mu_{1,e}(1-\rho'_{1,(i+1),e})} \tag{9}$$

Here, we have used the Laplace transform for the system time distribution in an M/M/1 queue. (With an arrival rate λ, a service rate μ and $\rho = \lambda/\mu$, it is given by $\frac{\mu(1-\rho)}{s+\mu(1-\rho)}$ [15, 16].)

Assume a packet requires k transmissions to be accepted at receiver L-layer without errors ($1 \le k \le M_1$). This happens with the probability $q_1 p_1^{k-1}$. In this case, during the first $k-1$ transmissions the packet goes through the sender L-layer queue $k-1$ times and the time-out $k-1$ times. At the last (successful) transmission, the packet goes through the sender L-layer queue, propagates along the link, and gets processed at the receiver L-layer queue. Since the time spent by a packet in the sender L-layer queue, a time-out period, the propagation delay on a link, and the time spend by a packet in the receiver L-layer queue are independent, the Laplace transform $\Phi^*_{1,i}(s)$ for the distribution of the sojourn time (time starting when a packet first enters the sender L-layer queue at node i and ending when the packet is passed to the PN-layer queue at node $i+1$, including the time incurred by retransmissions) becomes

$$\begin{aligned}\Phi^*_{1,i}(s) &= \sum_{k=1}^{M_1}[q_1 p_1^{k-1}\{F^*_{1,i,t}(s)S^*_1(s)\}^{k-1}F^*_{1,i,t}(s)G^*_1(s)F^*_{1,(i+1),e}(s)]\frac{1}{\sum_{k=1}^{M_1} q_1 p_1^{k-1}} \\ &= \sum_{k=1}^{M_1}\{q_1 p_1^{k-1}S^*_1(s)^{k-1}F^*_{1,i,t}(s)^k G^*_1(s)F^*_{1,(i+1),e}(s)\}\frac{1}{1-p_1^{M_1}} \\ &= \frac{1}{1-p_1^{M_1}}G^*_1(s)F^*_{1,(i+1),e}(s)A^*_1(s) \end{aligned} \tag{10}$$

where $A^*_1(s) = \sum_{k=1}^{M_1}\{q_1 p_1^{k-1}S^*_1(s)^{k-1}F^*_{1,i,t}(s)^k\}$; $S^*_1(s)$ is the Laplace transform for the link-by-link time-out period ($S^*_1(s) = e^{-s\tau_1}$); $G^*_1(s)$ is the Laplace transform for the propagation delay ($G^*_1(s) = e^{-sD_{prop}}$); and $\frac{1}{1-p_1^{M_1}}(= \frac{1}{\sum_{k=1}^{M_1} q_1 p_1^{k-1}})$ is a normalizing factor. Note that delays only for successful packets are considered. Note also that the time-out period and the propagation delay between two adjacent nodes are constant.

$F^*_{3,t}(s)$ and $F^*_{2,i}(s)$ in Eq.(7) are easily obtained in the following way. $F^*_{3,t}(s)$ is the Laplace transform of the distribution of the time spent by a packet in a source PT-layer queue, and $F^*_{2,i}(s)$ is the Laplace transform of the distribution of the time spent by a packet in a PN-layer queue at node i. From the same argument used to obtain $F^*_{1,i,t}(s)$, we have

$$F^*_{3,t}(s) = \frac{\mu_{3,t}(1-\rho_{3,t})}{s+\mu_{3,t}(1-\rho_{3,t})} \tag{11}$$

$$F^*_{2,i}(s) = \frac{\mu_2(1-\rho_{2,i})}{s+\mu_2(1-\rho_{2,i})} \tag{12}$$

By substituting Eqs.(10), (11) and (12) into Eq.(7), we can obtain $B^*_3(s)$, the Laplace transform for the distribution of the end-to-end packet transfer delay. From this Laplace transform, the average T, the second moment $T^{(2)}$, and the standard deviation T_σ of the end-to-end transfer delay of packets are obtained as follows:

$$T = -\frac{d}{ds}B^*_3(s)|_{s=0} \tag{13}$$

$$T^{(2)} = \frac{d^2}{ds^2}B^*_3(s)|_{s=0} \tag{14}$$

$$T_\sigma = \sqrt{T^{(2)} - T^2} \tag{15}$$

In the following, we define

$$\Theta' = \frac{d}{ds}\Theta^*(s)|_{s=0}, \qquad \Theta'' = \frac{d^2}{ds^2}\Theta^*(s)|_{s=0} \tag{16}$$

where $\Theta^*(s)$ is a Laplace transform of a distribution. From simple manipulation, T and $T^{(2)}$ become

$$T = -B_3' = -F_{3,t}' - F_{2,1}' - \sum_{i=1}^{l}(\Phi_{1,i}' + F_{2,(i+1)}')) \tag{17}$$

$$\begin{aligned} T^{(2)} &= B_3'' \\ &= F_{3,t}'' + 2F_{3,t}'F_{2,1}' + F_{2,1}'' + 2(F_{3,t}' + F_2')\sum_{i=1}^{l}(\Phi_{1,i}' + F_{2,(i+1)}') \\ &\quad + \sum_{i=1}^{l}\{\Phi_{1,i}'' + F_{2,(i+1)}'' - (\Phi_{1,i}')^2 - (F_{2,(i+1)}')^2\} + \{\sum_{i=1}^{l}(\Phi_{1,i}' + F_{2,(i+1)}')\}^2 \end{aligned} \tag{18}$$

where

$$\Phi_{1,i}' = G_1' + F_{1,(i+1),e}' + \frac{A_1'}{1 - p_1^{M_1}} \tag{19}$$

$$\Phi_{1,i}'' = G_1'' + F_{1,(i+1),e}'' + 2G_1'F_{1,(i+1),e}' + \frac{1}{1 - p_1^{M_1}}\{A_1'' + 2A_1'(G_1' + F_{1,(i+1),e}')\} \tag{20}$$

$$A_1' = \frac{(M_1 - 1)p_1^{M_1+1} - M_1 p_1^{M_1} + p_1}{q_1}S_1' + \frac{1 - (M_1+1)p_1^{M_1} + M_1 p_1^{M_1+1}}{q_1}F_{1,i,t}' \tag{21}$$

$$\begin{aligned} A_1'' =& \frac{2M_1(M_1-2)p_1^{M_1-1} - M_1(M_1-1)p_1^{M_1-2} - (M_1-1)(M_1-2)p_1^{M_1} + 2}{q_1^2}(S_1')^2 p_1^2 \\ &+ \frac{p_1\{-M_1(M_1+1)p_1^{M_1-1} + 2(M_1+1)(M_1-1)p_1^{M_1} - M_1(M_1-1)p_1^{M_1+1} + 2\}}{q_1^2}\{2S_1'F_{1,i,t}' + (F_{1,i,t}')^2\} \\ &+ \frac{1 + (M_1-1)p_1^{M_1} - M_1 p_1^{M_1-1}}{q_1}S_1'' p_1 + \frac{1 - (M_1+1)p_1^{M_1} + M_1 p_1^{M_1+1}}{q_1}F_{1,i,t}'' \end{aligned} \tag{22}$$

$$F_{1,i,t}' = -\frac{1}{\mu_{1,t}(1 - \rho_{1,i,t}')}, \qquad F_{1,i,e}' = -\frac{1}{\mu_{1,e}(1 - \rho_{1,i,e}')} \tag{23}$$

$$F_{2,i}' = -\frac{1}{\mu_2(1 - \rho_{2,i})}, \qquad F_{3,t}' = -\frac{1}{\mu_{3,t}(1 - \rho_{3,t})} \tag{24}$$

$$F_{1,i,t}'' = \frac{2}{\mu_{1,t}^2(1 - \rho_{1,i,t}')^2}, \qquad F_{1,i,e}'' = \frac{2}{\mu_{1,e}^2(1 - \rho_{1,i,e}')^2} \tag{25}$$

$$F''_{2,i} = \frac{2}{\mu_2^2(1-\rho_{2,i})^2}, \qquad F''_{3,t} = \frac{2}{\mu_{3,t}^2(1-\rho_{3,t})^2} \tag{26}$$

$$S'_1 = -\tau_1, \qquad G'_1 = -D_{prop}, \qquad S''_1 = \tau_1^2, \qquad G''_1 = D_{prop}^2. \tag{27}$$

At this point, the unknown factor in Eqs.(13), (14) and (15) is only $r_{1,i}$. Since $r_{1,i}$ is the average time interval from a packet's arrival at the sender L-layer queue to the end of link-by-link time-out period, $r_{1,i} = -F'_{1,i,t} + \tau_1$. (Refer to the definition of $r_{1,i}$ in subsection 3.3 and also to Figure 2(b).)

4.3 Edge-to-Edge Scheme

In this section, we focus on an edge-to-edge error recovery scheme and analyze its performance.

Effective Bit Rate

First, we obtain the effective packet arrival rate λ'_3 (i.e., the rate of the aggregated arrivals of the new packets and the retransmitted packets) at the PT-layer queue of the source node (see Figure 1). From the same argument used to obtain the effective packet arrival rate $\lambda'_{1,i}$ at the L-layer queue at node i in a link-by-link scheme (see Eq.(5)), we have

$$\lambda'_3 = \lambda_3[N_3 \frac{p_3\{1 - M_3 p_3^{M_3-1} + (M_3-1)p_3^{M_3}\}}{(1-p_3)^2} + \frac{1-p_3^{M_3}}{1-p_3}]. \tag{28}$$

where $N_3 = r_3\lambda_3$ for Go-Back-N, and $N_3 = 0$ for Selective-Repeat.

At the source node, since the output from the PT-layer queue is the input to its PN-layer queue and in turn the input to its L-layer queue, we have $\lambda'_3 = \lambda_2 = \lambda_1$. Note that with an edge-to-edge scheme, no packet will be dropped at the intermediate nodes, and thus the packet arrival rate at intermediate switching nodes does not change along the path. Packets are dropped at the destination PT-layer only if they have errors after M_3 edge-to-edge retransmissions. Therefore, the effective utilization of each queue becomes

- $\rho_{1,t} = \lambda_1/\mu_{1,t}$ at a sender L-layer queue,
- $\rho_2 = \lambda_2/\mu_2$ at a sender PN-layer queue,
- $\rho'_{3,t} = \lambda'_3/\mu_{3,t}$ at the source PT-layer queue, and
- $\rho'_{3,e} = \lambda'_3/\mu_{3,e}$ at the destination PT-layer queue.

The utilization of a receiver L-layer queue $\rho_{1,e}$ at intermediate switching nodes is not considered since no processing is required at a receiver L-layer queue.

End-to-End Packet Transfer Delay

An argument similar to the one used in the link-by-link scheme applies to obtain the Laplace transform $B_3^*(s)$ for the distribution of the end-to-end packet transfer delay in the edge-to-edge scheme.

Assume a packet, say a test packet, requires k edge-to-edge transmissions to be accepted at the destination PT-layer without errors ($1 \leq k \leq M_3$). This happens with the probability

$q_3 p_3^{k-1}$ where $q_3 = 1 - p_3$. Recall that p_3 is the probability that a packet arrives at the destination PT-layer with an error (or errors).

During the first $k-1$ retransmissions, the test packet goes through the PT-layer queue at the source node $k-1$ times and an edge-to-edge time-out $k-1$ times. (Note that an edge-to-edge time-out starts right after the source PT-layer queue passes a packet to the PN-layer queue.) The Laplace transform for the distribution of the time spent for these $k-1$ retransmissions is $\{F_{3,t}^*(s)S_3^*(s)\}^{k-1}$. Here, $F_{3,t}^*(s)$ is the Laplace transform for the distribution of the time spent at the source PT-layer queue. Since the input to the PT-layer queue at the source node follows a Poisson process, $F_{3,t}^*(s)$ is easily obtained by using the corresponding ρ value in the Laplace transform for the system time distribution in an M/M/1 queue. $S_3^*(s)$ is the Laplace transform of the edge-to-edge time-out period, and is given by $S_3^*(s) = e^{-s\tau_3}$.

For the last (successful) transmission, the test packet experiences the following delays: (1) the time spent at the PT-layer queue at the source, (2) the time spent at the PN-layer queue at the source, (3) the time spent at intermediate switching nodes (4) the time spent at the PN-layer queue at the destination, and (5) the time spent at the PT-layer queue at the destination. The Laplace transform for the delay element (1) is $F_{3,t}^*(s)$. The Laplace transform $F_2^*(s)$ for the delay element (2) is easily obtained from the Laplace transform for the system time distribution in an M/M/1 queue since the input to the PN-layer queue at the source node follows a Poisson process. Assuming that there are l hops from the source to the destination, the Laplace transform for (3) and (4) combined is given by $\{F_{1,t}^*(s)G_1^*(s)F_2^*(s)\}^l$, where $F_{1,t}^*(s)$ is the Laplace transform for the time spent at a sender L-layer queue (and is obtained by using a corresponding ρ value in the Laplace transform for the system time distribution in an M/M/1 queue); $G_1^*(s)$ is the Laplace transform for the propagation delay ($G_1^*(s) = e^{-sD_{prop}}$); and $F_2^*(s)$ is the Laplace transform for the time spent at a PN-layer queue. Note that the time spent at the L-layer queue at a receiving node ($F_{1,e}^*(s)$) is not included since no processing is required at this layer in an edge-to-edge scheme. The Laplace transform $F_{3,e}^*(s)$ for the delay element (5) is obtained from the Laplace transform for the system time distribution in an M/M/1 queue, since the input to the PT-layer queue at the destination node follows a Poisson process.

From the above discussion, $B_3^*(s)$ is given by

$$\begin{aligned}
B_3^*(s) &= \sum_{k=1}^{M_3}[q_3 p_3^{k-1}\{F_{3,t}^*(s)S_3^*(s)\}^{k-1}F_{3,t}^*(s)F_2^*(s)\{F_{1,t}^*(s)G_1^*(s)F_2^*(s)\}^l F_{3,e}^*(s)]\frac{1}{\sum_{k=1}^{M_3} q_3 p_3^{k-1}} \\
&= \sum_{k=1}^{M_3}\{q_3 p_3^{k-1} S_3^*(s)^{k-1}F_{3,t}^*(s)^k F_2^*(s)^{l+1}F_{1,t}^*(s)^l G_1^*(s)^l F_{3,e}^*(s)\}\frac{1}{1-p_3^{M_3}} \\
&= \frac{1}{1-p_3^{M_3}}F_2^*(s)^{l+1}F_{1,t}^*(s)^l G_1^*(s)^l F_{3,e}^*(s)A_3^*(s) \qquad (29)
\end{aligned}$$

where $A_3^*(s) = \sum_{k=1}^{M_3}\{q_3 p_3^{k-1} S_3^*(s)^{k-1}F_{3,t}^*(s)^k\}$; and $\frac{1}{1-p_3^{M_3}} (= \frac{1}{\sum_{k=1}^{M_3} q_3 p_3^{k-1}})$ is a normalizing factor.

From Eq.(29), the average T and the second moment $T^{(2)}$ of the end-to-end packet transfer delay become

$$T = -B_3' = -(l+1)F_2' - lF_{1,t}' - lG_1' - F_{3,e}' - \frac{A_3'}{1-p_3^{M_3}} \tag{30}$$

$$\begin{aligned}
T^{(2)} &= B_3'' \\
&= (l+1)F_2'' + lF_{1,t}'' + lG_1'' + F_{3,e}'' + l(l+1)(F_2')^2 + l(l-1)(F_{1,t}')^2 + l(l-1)(G_1')^2 \\
&\quad + 2l(l+1)F_2'F_{1,t}' + 2l(l+1)F_2'G_1' + 2(l+1)F_2'F_{3,e}' + 2l^2F_{1,t}'G_1' + 2lG_1'F_{3,e}' + 2lF_{1,t}'F_{3,e}' \\
&\quad + \frac{2A_3'\{(l+1)F_2' + lF_{1,t}' + lG_1' + F_{3,e}'\}}{1-p_3^{M_3}} + \frac{A_3''}{1-p_3^{M_3}}
\end{aligned} \tag{31}$$

where $S_3' = -\tau_3$ and $S_3'' = \tau_3^2$. The rest of the parameters in Eqs.(30) and (31) can be easily obtained using the methods used in section 4.2. r_3 is given by $-F_{3,t}' + \tau_3$ (see Figure 2(a)).

5. Numerical Examples

In this section we show some numerical examples for the link-by-link and the edge-to-edge error recovery schemes. In the figures, the following notations are used to represent error recovery schemes:

- E-to-E : edge-to-edge error recovery scheme, and
- L-by-L : link-by-link error recovery scheme.

5.1 Packet Loss Probability

The packet loss probability ε across a network (i.e., the probability that a packet is discarded due to the limitation on the maximum number of retransmissions allowed) is a function of M_1 (the maximum number of packet transmissions allowed at L-layer) in a case where a link-by-link error recovery scheme is used, or a function of M_3 (the maximum number of packet transmissions allowed at PT-layer) in a case where an edge-to-edge scheme is used.

Figure 6 shows the effect of the values of M_1 and M_3 on the packet loss probability across a network when 500 bit-long packets are transferred through 4 hops. A packet size of 500 bits is used since it is close to the CCITT standard ATM cell size of 424 bits (i.e., 53 octets) [17, 18]. The horizontal axis shows the packet error rate on a link (i.e., the probability that a packet receives an error on a link). Note that it is not a "bit" error rate. M_1 and M_3 are the parameters, and their values are indicated by a tuple (M_1, M_3) in this figure. It is apparent that "no error recovery scheme ($M_1 = M_3 = 1$)" gives the worst loss probability. The loss probability improves with the increase in the values of M_1 and M_3. For instance, when the packet error rate on a link p_1 is 10^{-6}, which corresponds to 2×10^{-9} bit error rate assuming randomly distributed errors, a maximum of two transmissions for both a link-by-link scheme (indicated by the line (2, 1) in Figure 6) and an edge-to-edge scheme (indicated by the line (1, 2)) gives less than 10^{-10} packet loss probability across a network.

Since optical fibers can easily achieve a packet error rate of 10^{-6} (or equivalently, a bit error rate of 2×10^{-9}) on a link, and since 10^{-10} packet loss probability across the network is small enough to satisfy the requirements for broadband networks (i.e., the cross-network packet loss rate of less than 10^{-9}), in most of the following numerical examples it is assumed that $M_1 = M_3 = 2$ and $p_1 = 10^{-6}$.

5.2 Packet Transfer Delay

Throughout the numerical examples in this subsection, unless otherwise stated, we assume the following parameter values: the average packet length P = 500 bits, the channel speed V = 150 Mbits/sec, the average packet transmission time $= 1/\mu_{1t} = P/V \approx 3.333 \times 10^{-6}$ seconds, the packet error rate on a link $p_1 = 10^{-6}$, and the maximum numbers of packet transmissions allowed at L-layer and at PT-layer $M_1 = M_3 = 2$. These parameter values are chosen such that they are very close to the actual parameter values in ATM networks. For instance, the packet size of 500 bits is very close to the CCITT standard ATM cell size of 424 bits, and the channel speed of 150 Mbits/sec is very close to the effective data rate of 149.76 Mbits/sec when STM-1 of SDH (Synchronous Digital Hierarchy) is used for ATM transmission [19]. The use of optical fibers as a communication media can easily provide 10^{-6} packet error rate on a link, and the justification for $M_1 = M_3 = 2$ is discussed in the above subsection. It is also assumed that the number of hops l from a source to a destination is 4, and a link propagation delay D_{prop} is 30 times the average packet transmission time (i.e., an internode distance of 20 Km). The average processing time at each layer is assumed to be comparative to the average packet transmission time $\frac{1}{\mu_{1t}}$. Namely, we assume that $\frac{1}{\mu_{1e}} = \frac{1}{\mu_2} = \frac{1}{\mu_{3t}} = \frac{1}{\mu_{3e}} = \frac{1}{\mu_{1t}}$. The link-by-link time-out period τ_1 is set to 84 times the average packet transmission time. This time out period is the same value used in the simulations in subsection 4.1 to verify the accuracy of the Poisson assumption. The edge-to-edge time-out period τ_3 is set to 336 times the average packet transmission time, four times the link-by-link time-out period. Throughout the numerical examples in this subsection, an average packet transmission time ($= 1/\mu_{1t} = P/V \approx 3.333 \times 10^{-6}$ seconds) is used as a unit time.

In Figure 7 through Figure 10, Go-Back-N retransmission procedure is used in both edge-to-edge and link-by-link error recovery schemes. Figure 7 compares the average end-to-end packet transfer delay T in the link-by-link and in the edge-to-edge error recovery schemes when $p_1 = 10^{-3}$. 10^{-3} packet error rate represents the error rate in existing packet switched networks. The horizontal axis shows the rate of the input traffic, i.e., the packet arrival rate λ_3 at PT-layer at the source node. In this figure, the edge-to-edge scheme provides the smaller average transfer delay when the input traffic is light. As the input traffic increases, however, the edge-to-edge scheme gives larger delay than the link-by-link scheme and reaches the network saturation point sooner than the link-by-link scheme. This is due to the trade-off between the protocol-processing overhead and the time to recover from the error. Because the edge-to-edge scheme eliminates hop-by-hop error checking and retransmissions, the protocol-processing overhead is smaller in the edge-to-edge scheme than in the link-by-link scheme. On the other hand, in the edge-to-edge scheme one retransmission requires larger delay since erred packets are retransmitted from the source to the destination, not between two adjacent nodes as in link-by-link scheme. When the input traffic is small, and thus the number of erred packets is small, this drawback of edge-to-edge scheme (i.e., longer delay required for a retransmission) is outweighed by the benefit of the reduced processing overhead.

In Figure 8, the packet error rate on a link p_1 is decreased to 10^{-6}, which represents the packet error rate in high-speed network environments. In this figure, the edge-to-edge schemes provides the smaller delay for all ranges of the input traffic. This is due to the following reason: when the error rate is extremely small, the number of erred packets is also small even when the input traffic is high. Therefore, the drawback of the edge-to-edge scheme (i.e., longer time required for a retransmission) is outweighed by the reduced processing overhead, and thus the edge-to-edge scheme gives the smaller delay. From Figures 7 and 8 it can be concluded that as the packet error rate decreases, the edge-to-edge scheme provides the smaller average transfer delay than the link-by-link scheme for the wider traffic range.

Figure 9 illustrates an optimal error recovery scheme to give the smallest average transfer delay for a given packet error rate p_1 and the input traffic rate λ_3. It is assumed that $M_3 = M_1 = 2$. The area "[E-to-E]" shows the area where the edge-to-edge scheme provides the smaller delay. The area "[L-by-L]" shows the area where the link-by-link scheme provides the smaller delay. The area "saturation" shows the area where a network becomes saturated. For instance, if $p_1 = 10^{-3}$, the edge-to-edge scheme yields a smaller delay than the link-by-link scheme in the traffic range $0 < \lambda_3 < 0.3$. Figure 9 shows that as the packet error probability p_1 on a link decreases, the edge-to-edge scheme yields a smaller delay for the wider traffic range, and eventually, the edge-to-edge scheme becomes superior to the link-by-link scheme for all the traffic range. For instance, for the packet error probability of 10^{-6}, the edge-to-edge scheme almost always (except for the extremely high input traffic close to 1) yields a smaller delay than the link-by-link scheme. Since it is commonly assumed that broadband networks will provide very low bit error rates (i.e., 10^{-9}), we can conclude that the edge-to-edge scheme is superior to the link-by-link scheme in high-speed networks.

Figure 10 shows the effect of the decreased processing time on the optimal error recovery scheme. In this figure, the protocol-processing time is assumed to be almost negligible (i.e., 0.001). Note that in Figure 9, the protocol-processing time is assumed to be 1. By comparing Figures 9 and 10 it can be seen that when processing time is decreased, the area where the link-by-link scheme gives the smaller delay becomes wider. This is because the benefit of reduced protocol-processing overhead in the edge-to-edge scheme decreases when the processing time at each node becomes smaller.

In Figure 11 through Figure 13, Selective-Repeat retransmission procedure is assumed in both edge-to-edge and link-by-link schemes. Figure 11 compares the average packet transfer delay T in the link-by-link and in the edge-to-edge error recovery schemes. The packet error rate on a link p_1 is 10^{-3}. This figure shows a significant departure from what is observed in Figure 7 (Go-Back-N case). As seen in Figure 11, when Selective-Repeat procedure is used, even when the packet error probability is relatively high (10^{-3}), the edge-to-edge scheme provides the smaller average transfer delay for the entire traffic range than the link-by-link scheme. Unlike Go-Back-N procedure where a sender retransmits all the packets starting with the erred packet, in Selective-Repeat, only the erred packet is retransmitted. Therefore, even when the packet error rate is high, the number of retransmitted packets is small so that the advantage of the reduced processing overhead in the edge-to-edge scheme outweighs the drawback of longer time required for retransmissions.

In Figure 12, the packet error rate on a link p_1 is decreased to 10^{-6}. Figure 12 shows almost identical results to Figure 8 (Go-Back-N case). This is because the packet error probability used (10^{-6}) is too small to see the significant difference between Go-Back-N and Selective-Repeat procedures.

Figure 13 assumes Selective-Repeat retransmission procedure and illustrates an optimal error recovery scheme for a given packet error rate p_1 and the input traffic rate λ_3. By comparing this figure with Figure 9 (Go-Back-N case), it can be seen that with Selective-Repeat procedure, even when the packet error rate is relatively high, the edge-to-edge scheme provides a smaller delay than the link-by-link scheme for the wider traffic range. This result suggests that the superiority of the edge-to-edge scheme to the link-by-link scheme is more significant when Selective-Repeat is used.

Finally, the effect of the number of hops l between the source and the destination on the average transfer delay is examined. For each of the retransmission procedures, Table 3 shows the number of hops at which the performance of link-by-link surpasses that of edge-to-edge scheme (crossover point). The packet error probability on a link p_1 is the parameter in this table. The packet arrival rate λ_3 at PT-layer of the source node is fixed to 0.5. For instance, when p_1 is 10^{-6}, λ_3 is 0.5, and Go-Back-N procedure is used, the edge-to-edge scheme provides the smaller transfer delay than the link-by-link scheme for a network with less than 89 hops. For a network with 89 hops or more, the link-by-link scheme provides the smaller transfer delay. This result implies that as the number of hops increases, the performance of the link-by-link scheme approaches that of the edge-to-edge scheme, and eventually, the former will surpass the latter. This is because the inefficiency of the edge-to-edge scheme increases as the number of hops increases; as the number of hops increases, retransmissions between the source and the destination require longer time.

From Table 3, it can also be seen that the crossover in Selective-Repeat procedure happens at a larger value of l than in Go-Back-N procedure. This is because the number of packets needed to be retransmitted in Selective-Repeat is much smaller than that in Go-Back-N. Therefore, with Selective-Repeat, the edge-to-edge scheme performs better than the link-by-link scheme in the wider range of parameter values. Table 3 also indicates that as the error probability increases, the crossover happens at a smaller value of l. This is due to the following reason: as the error probability increases, more number of packets are retransmitted. Since one retransmission takes longer time in the edge-to-edge scheme than in the link-by-link scheme, the performance of the edge-to-edge scheme becomes significantly worse. Therefore, the crossover happens at a smaller value of l when the error probability increases.

6. Conclusion

This paper investigates an edge-to-edge error recovery scheme for a high speed packet switched network and obtains both the packet transfer delay and the packet loss probability across a network. The performance of an edge-to-edge scheme is compared with that of a link-by-link scheme, an error recovery scheme used in traditional networks. Through analysis, the effects of protocol-processing overhead on the performance of error recovery schemes are investigated. The effects of error probability and the number of hops between the source and the destination on the performance of the two error recovery schemes are also investigated. The results show that in high-speed network environments where the channel speed is very high and the error probability on a link is very small, the edge-to-edge scheme provides the smaller transfer delay than the link-by-link scheme for both both Go-back-N and Selective-Repeat retransmission procedures.

7. Acknowledgements

The authors would like to thank Mr. Nhuan Vu for his help in simulation programming in section 4.1.

References

[1] Bradley, Tracy T. and Suda, Tatsuya, "Survey of Unified Approaches to Integrated-Service Networks", Proc. of the IEEE International Telecommunications Symposium, Sept. 1990.

[2] Kulzer, J and Montgomery, W, "Statistical Switching Architecture for Future Services", ISS, 1984.

[3] Gonet, P, Adams, P and Coudreuse, J, "Asynchronous Time-Division Switching: The Way to Flexible Broadband Communication Networks", IEEE Int. Zurich Seminar on Digital Communications, Mar. 1986.

[4] Minzer, Steven E, "Broadband ISDN and Asynchronous Transfer Mode (ATM)", IEEE Communications Magazine, Vol.27, No.9, Sep. 1989.

[5] Hoberecht, W. L. "A Layered Network Protocol for Packet Voice and Data Integration", IEEE Journal on SAC, Vol.SAC-1, No.6, Dec. 1983.

[6] Murata, M. and Takagi, H., "Two-Layer Modeling for Local Area Networks", INFOCOM, 1987.

[7] Pennotti, M. C. and Schwartz, M., "Congestion Control in Store and Forward Tandem Links", IEEE Trans. Commun., Vol.COM-23, Dec. 1975.

[8] Bhargava, A., Kurose, J. F., Towsley, D. and Vanleemput, G., "Performance Comparison of Error Control Schemes in High-Speed Computer Communication Networks", IEEE Journal on SAC, Vol.6, No.9, Dec. 1988.

[9] Lam, S., "Store and Forward Buffer Requirements in a Packet Switching Network", IEEE Trans. Commun., Vol.COM-24, Apr. 1976.

[10] Irland, M. and Pujolle, G., "Comparison of Two Packet Retransmission Techniques", IEEE Trans. Inform. Theory, Vol.IT-26.

[11] Kuhl, D., "Error Recovery Protocols: Link-by-Link versus Edge-to-Edge", INFOCOM, 1983.

[12] Kleinrock, L., "Queueing Systems Volume II", John Wiley, 1976.

[13] Allen, A. O., "Probability, Statistics, and Queueing Theory", Academic Press, 1978.

[14] Jackson, J. R., "Jobshop-like Queueing Systems", Management Science 10, 1963.

[15] Sauer, C. H. and Chandy, K. M., "Computer Systems Performance Modeling", Prentice-hall, 1981.

[16] Wong, J. W., "Queueing Network Modeling of Computer Communication Networks", ACM Computing Surveys, Vol.10, No.3, Sept. 1978.

[17] CCITT Study Group XVIII, "WP XVIII/8–Report of Meeting", Temporary Document No.16 (Plenary), Geneva, Switzerland, Jun. 1989.

[18] CCITT Study Group XVIII, "Meeting Report of Sub-Working Party 8/1 ATM", Tem-

porary Document No.14-E (Plenary), Geneva, Switzerland, Jun. 1989.

[19] CCITT Recommendations, G707, G708 and G709, 1988.

	Layer	*Functions*	*Error Recovery Schemes*	
			E-to-E	L-by-L
Layer 3	PT-layer	End-to-End Packet Error Recovery	yes	no
	PN-layer	Routing	yes	yes
Layer 2	L-layer	Hop-by-Hop Frame Error Recovery	no	yes
		Frame Transmission	yes	yes

E-to-E: Edge-to-Edge Error Recovery Scheme
L-by-L: Link-by-Link Error Recovery Scheme

Table 1 Protocol Models

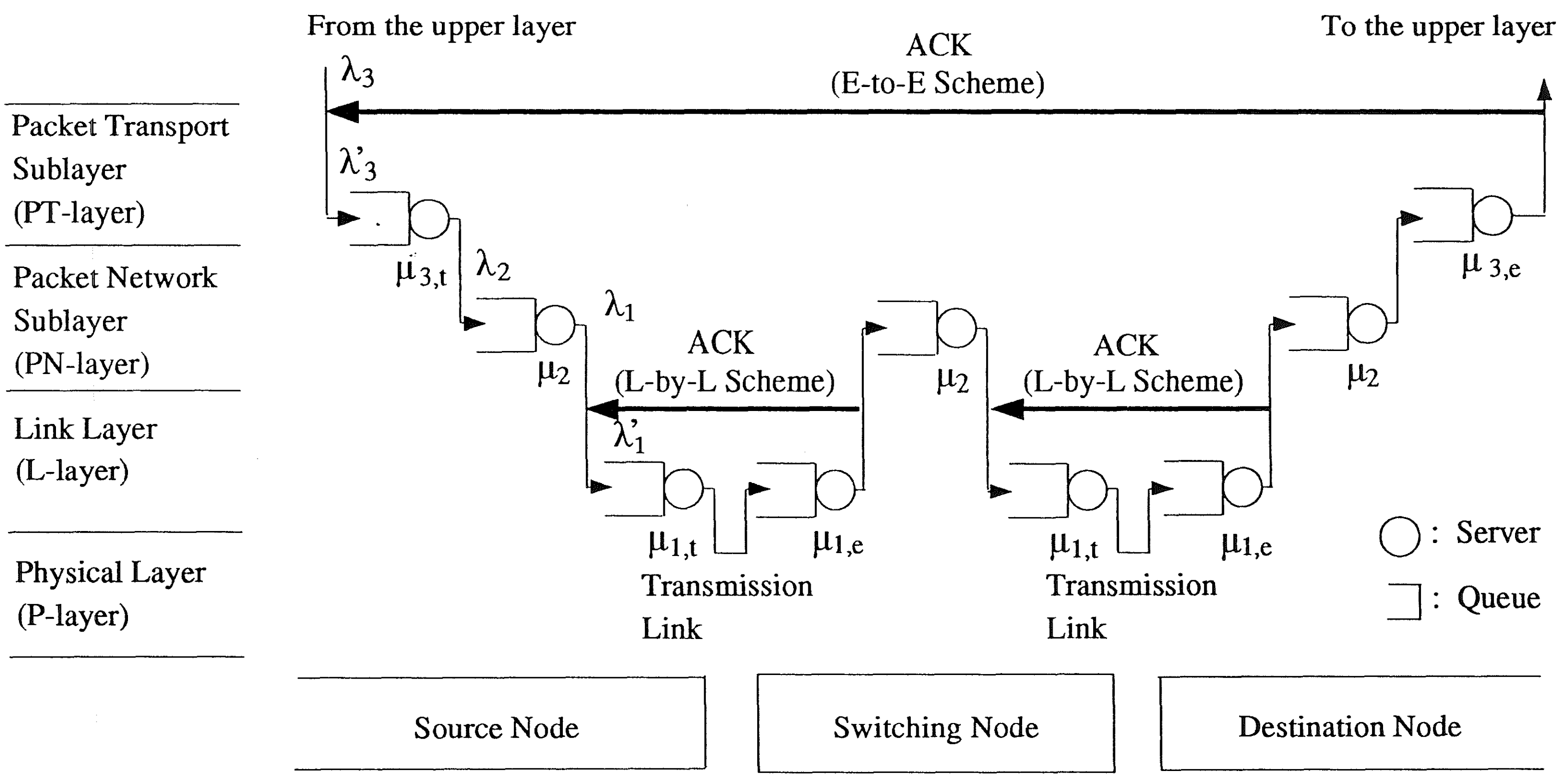
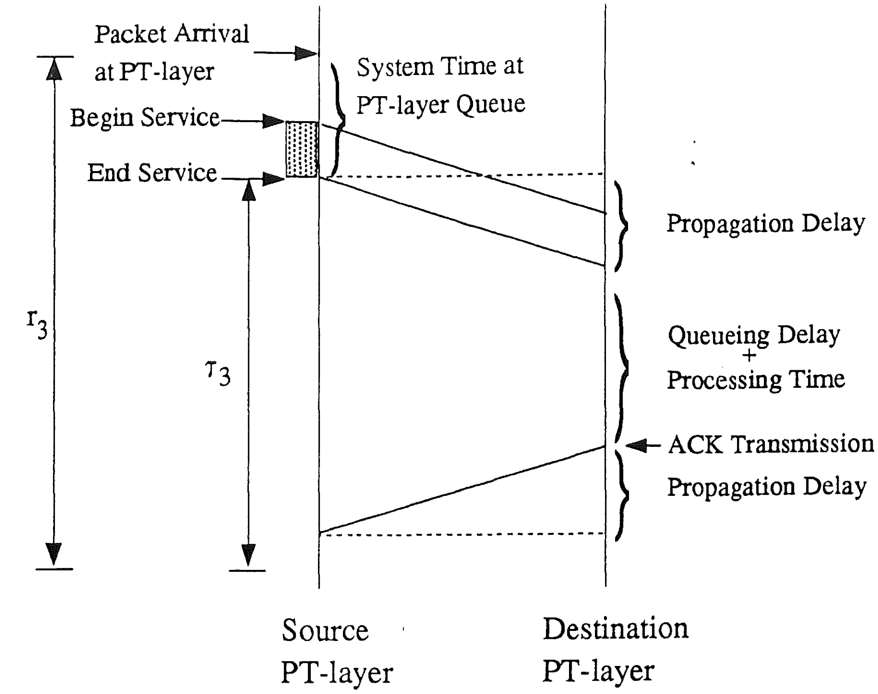
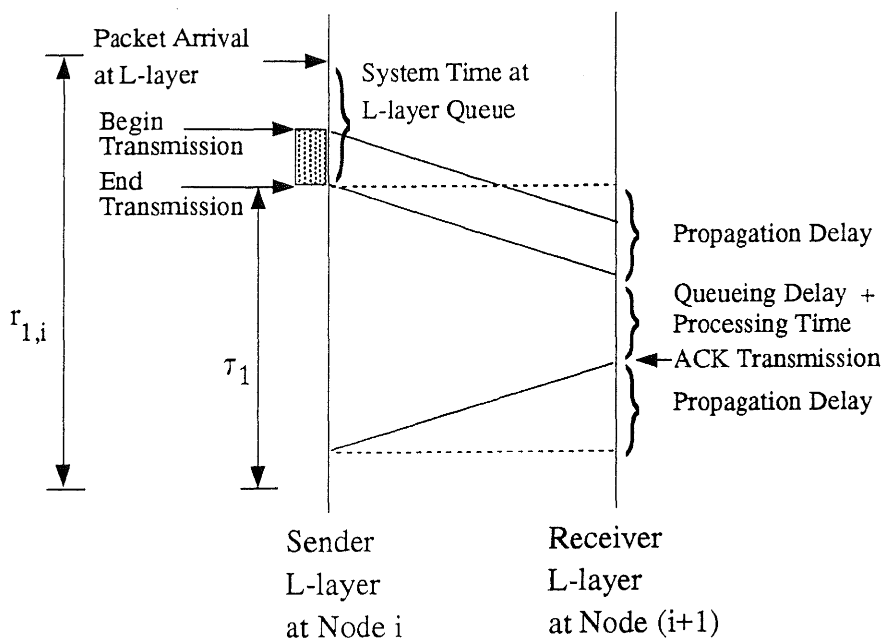


Figure 1 Queueing Network Model for Link-by-Link and Edge-to-Edge Error Recovery Schemes



(a) Time-out Period in the Edge-to-Edge Scheme



(b) Time-out Period in the Link-by-Link Scheme

Figure 2 Time-outs

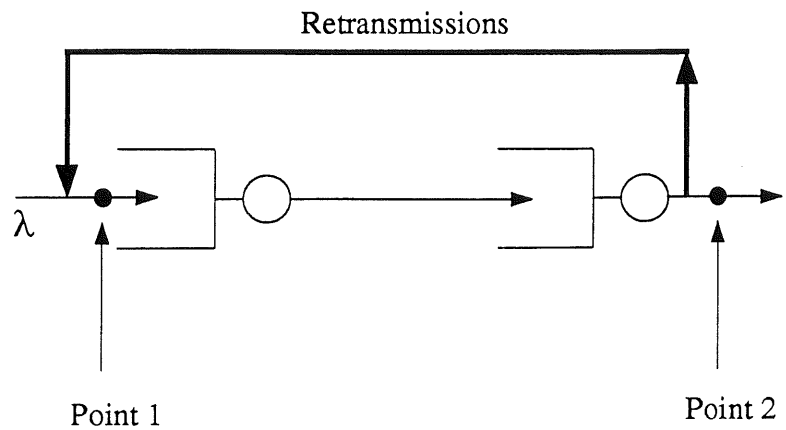


Figure 3 Simulation Model

Test Point / λ	Point 1	Point 2
0.3	Poisson	Poisson
0.5	Poisson	Poisson
0.7	Poisson	Poisson

Level of significance = 5%
Bit error rate = 2×10^{-7}

Table 2 Results of Chi-Square Goodness of Fit Test

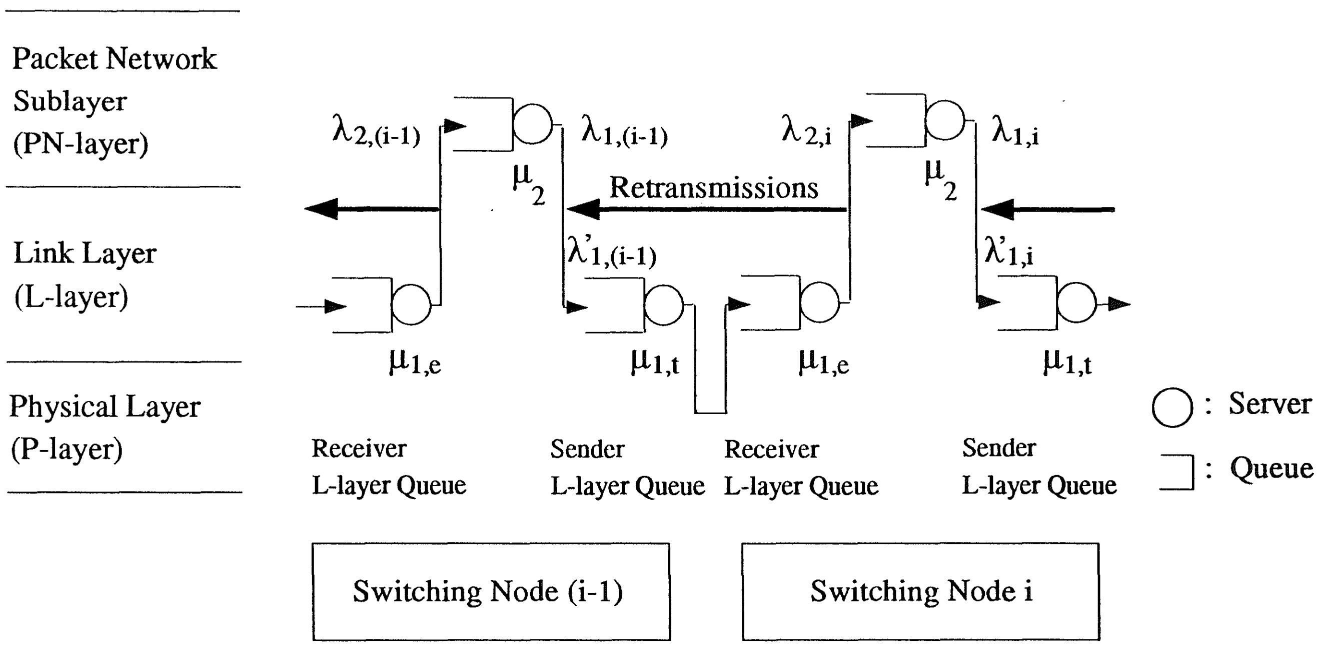


Figure 4 Arrival Rates at PN-layer and L-layer Queues

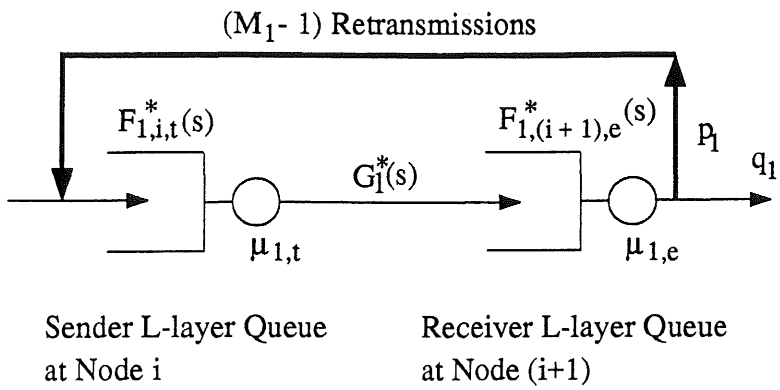


Figure 5 L-layer Queues

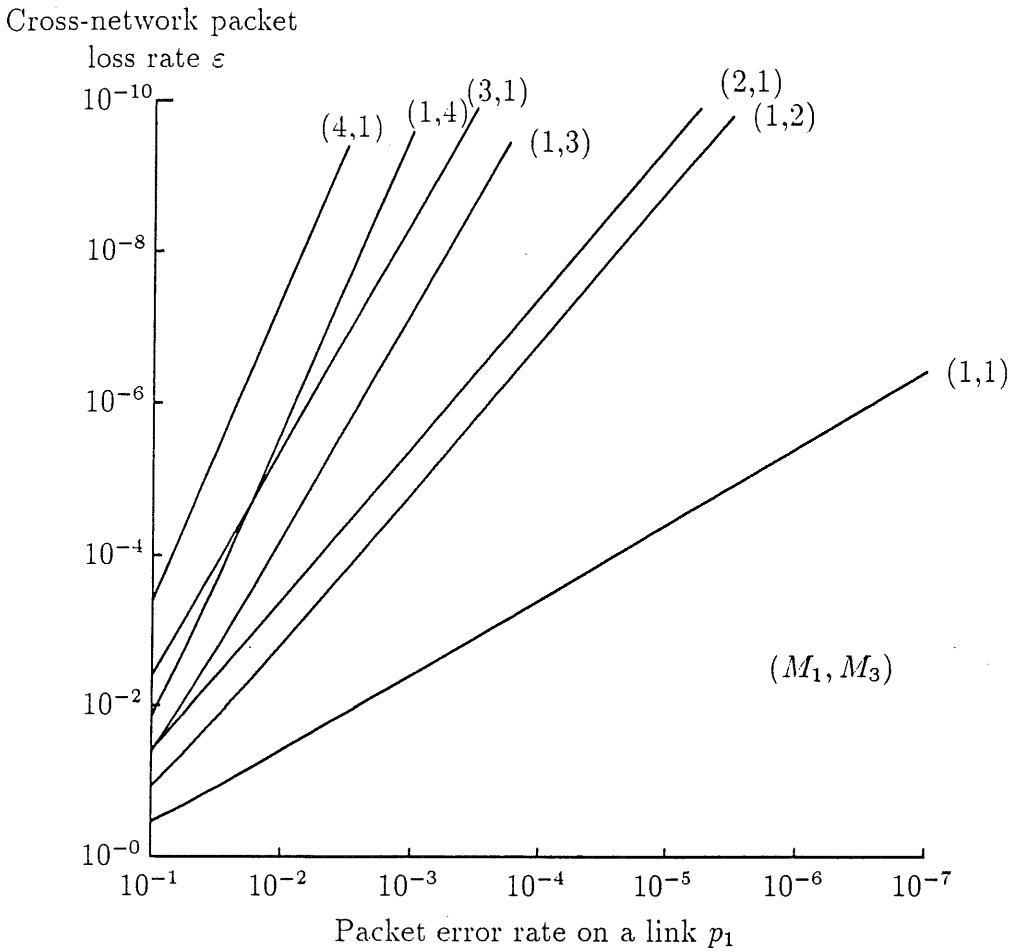


Figure 6 Cross-Network Packet Loss Rate
(4 hops)

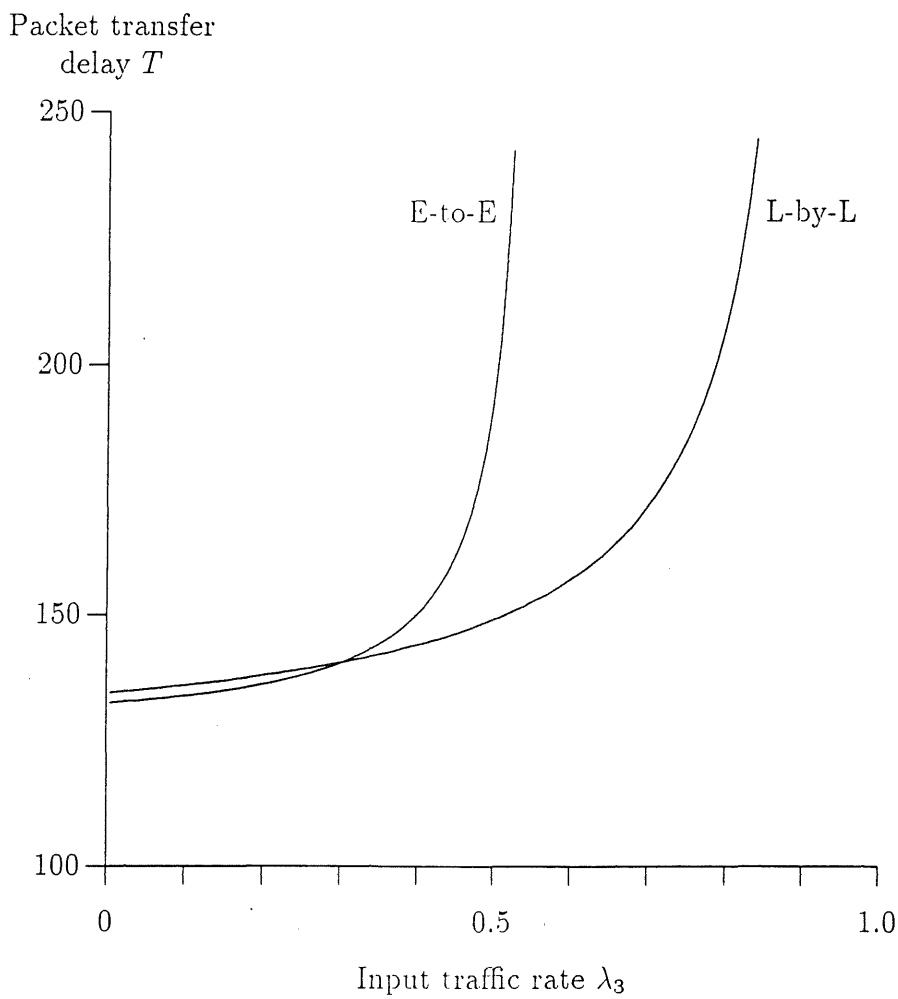


Figure 7 Average End-to-End Packet Transfer Delay Go-Back-N, 4 hops, $p_1 = 10^{-3}$

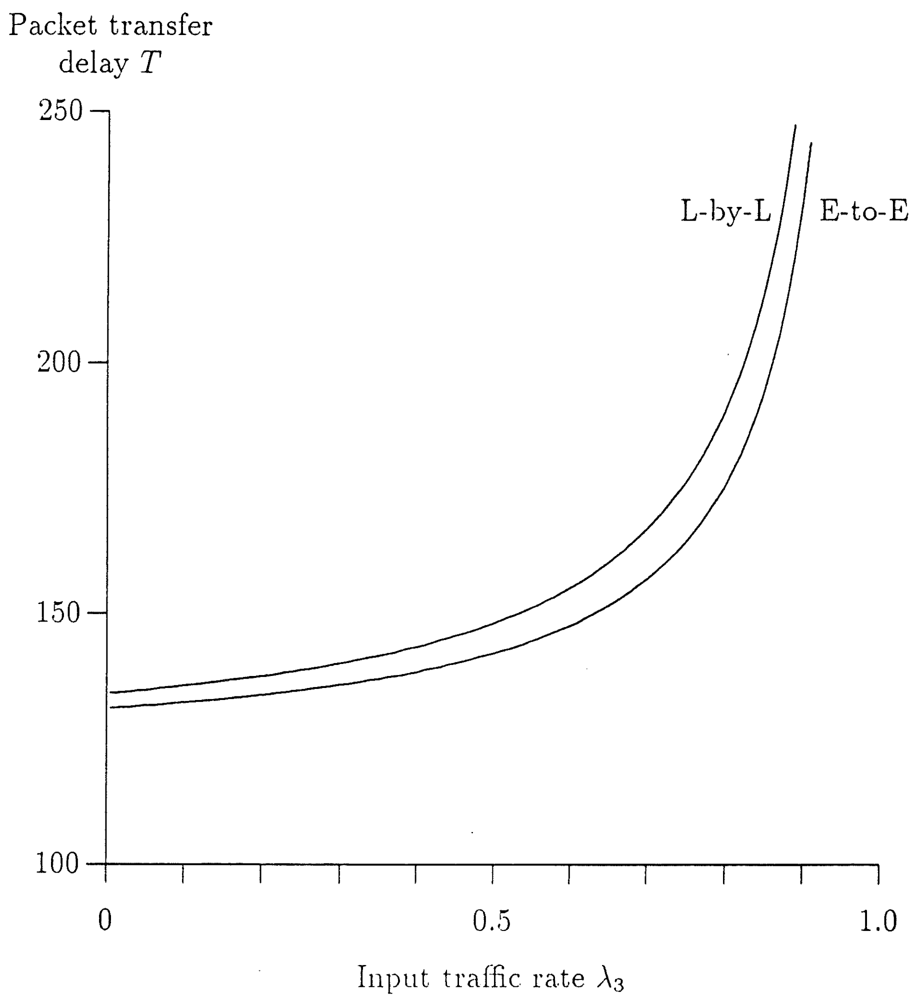


Figure 8 Average End-to-End Packet Transfer Delay Go-Back-N, 4 hops, $p_1 = 10^{-6}$

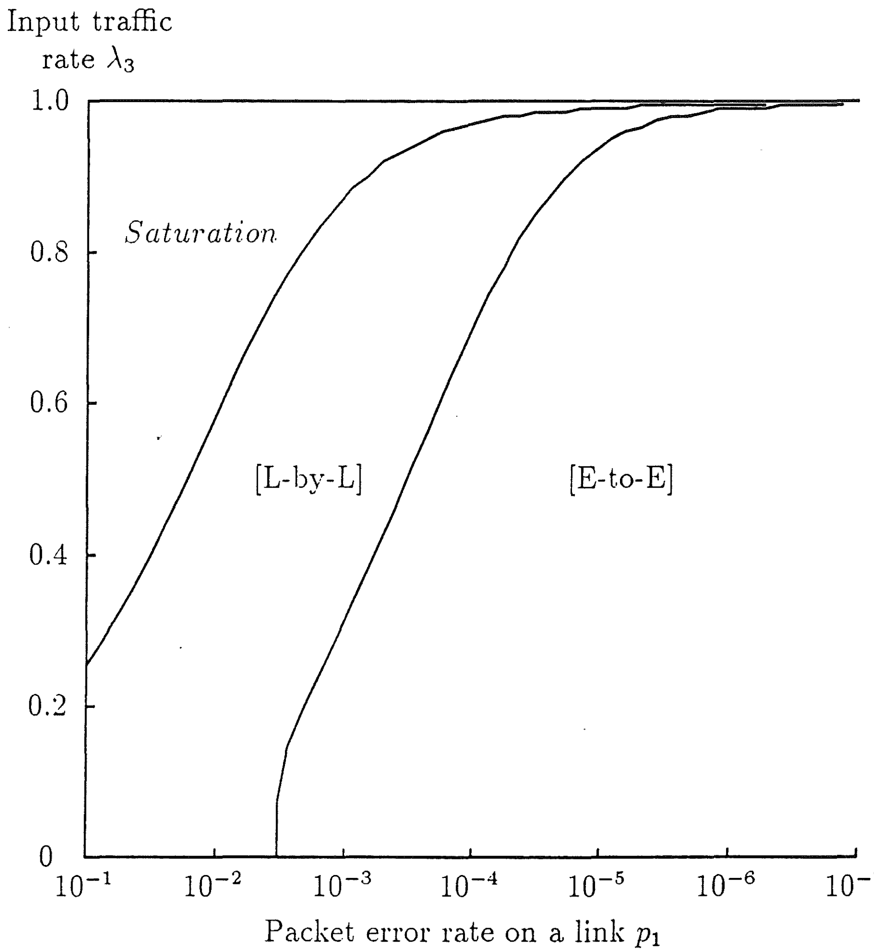


Figure 9 Optimal Error Recovery Scheme
Go-Back-N, 4 hops

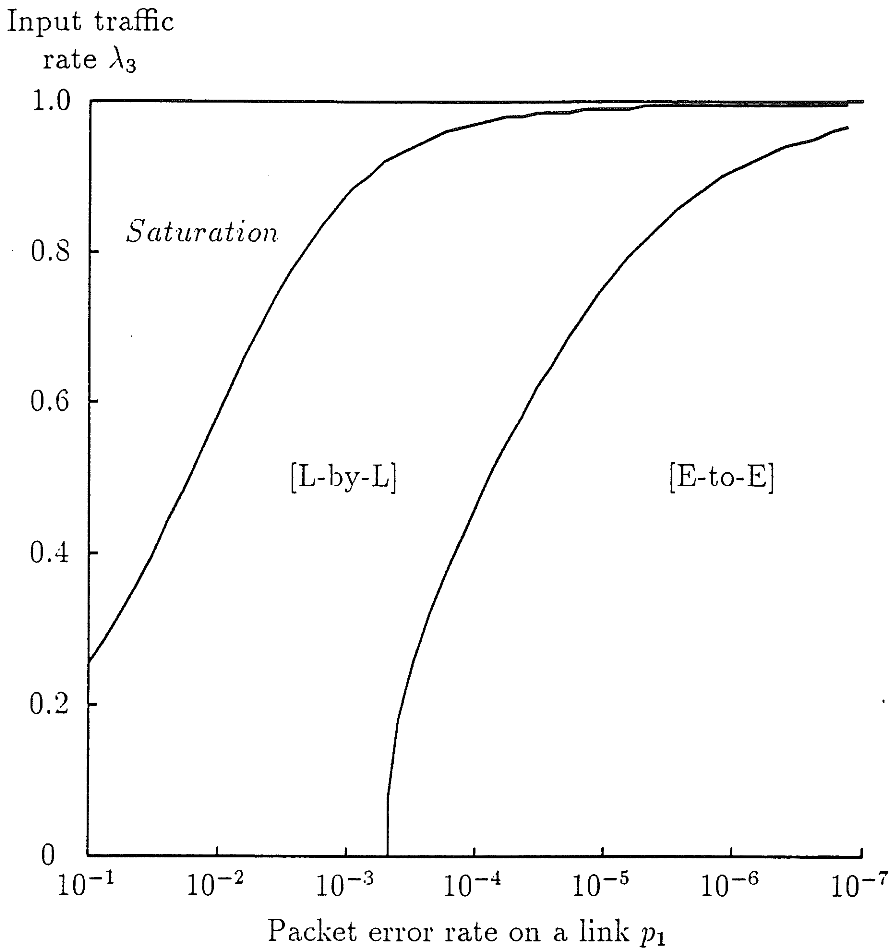


Figure 10 Effect of a Decreased Processing Time on Optimal Error Recovery Scheme Go-Back-N, 4 hops

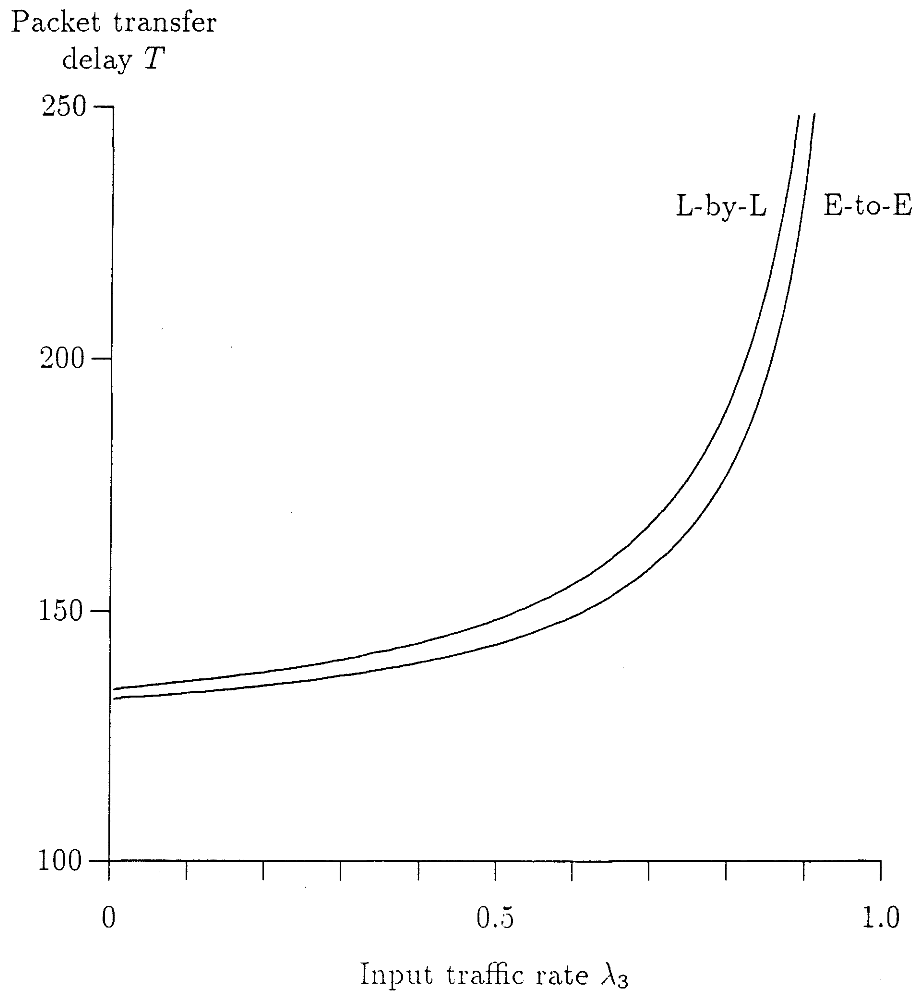


Figure 11 Average End-to-End Packet Transfer Delay
Selective-Repeat, 4 hops, $p_1 = 10^{-3}$

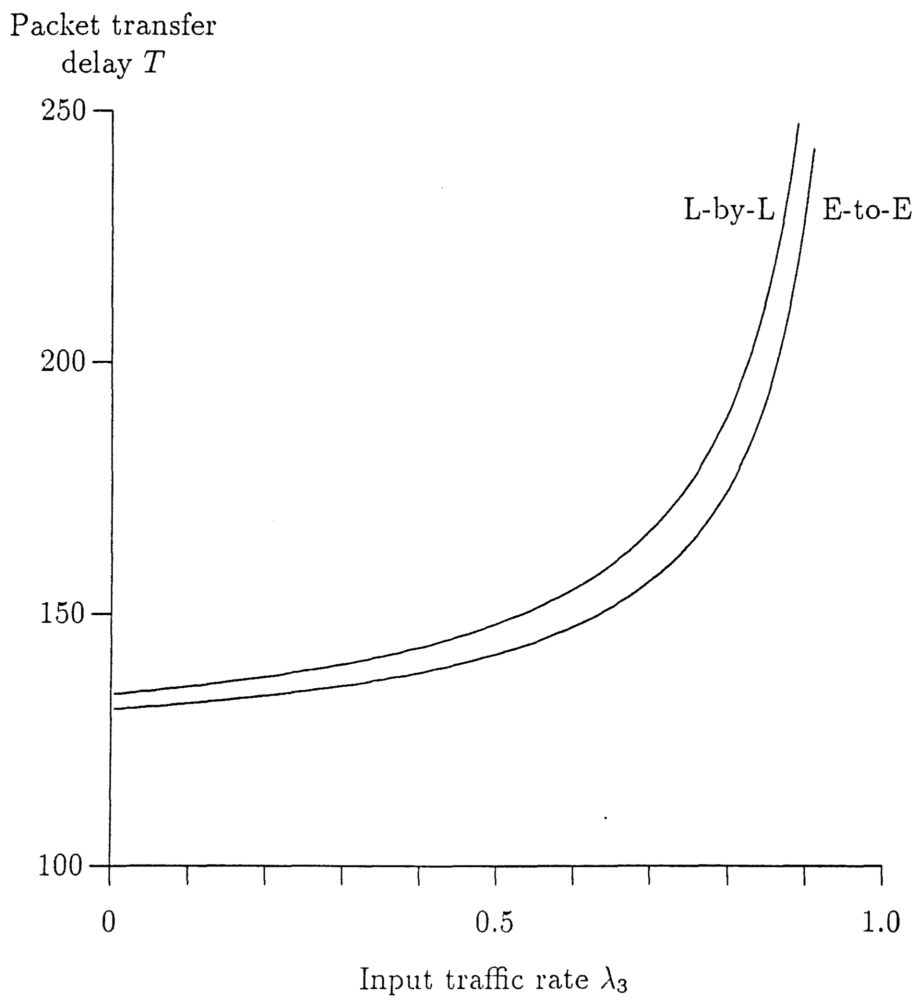


Figure 12 Average End-to-End Packet Transfer Delay Selective-Repeat, 4 hops, $p_1 = 10^{-6}$

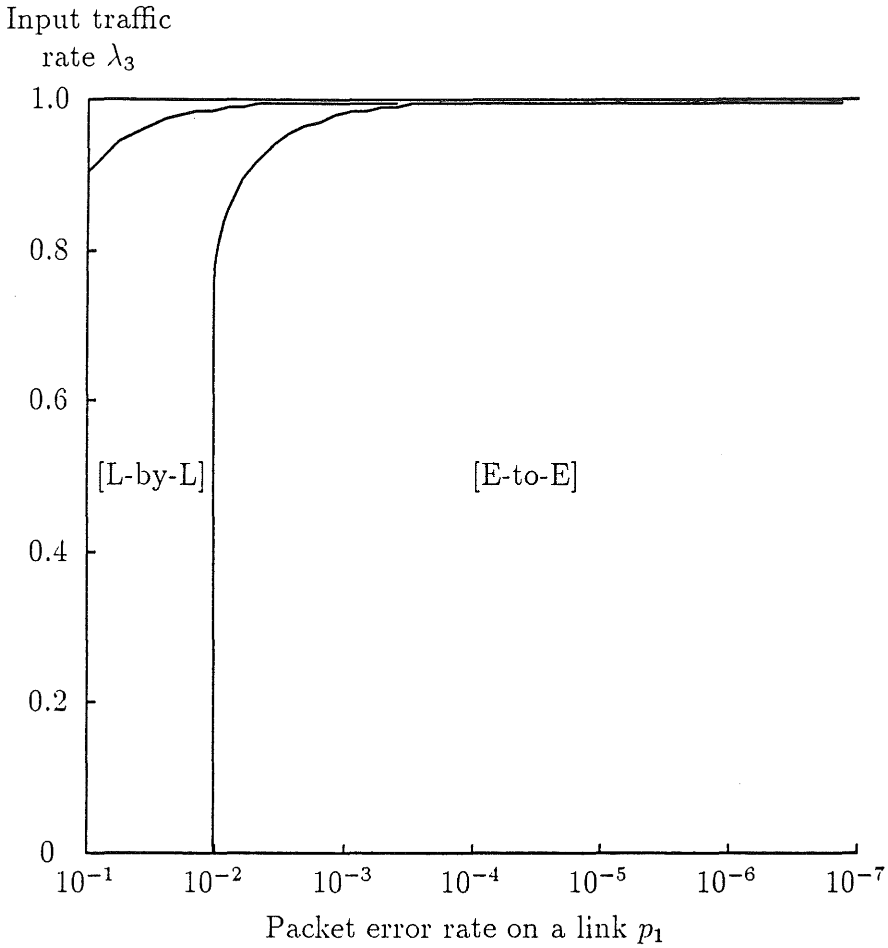


Figure 13 Optimal Error Recovery Scheme
Selective-Repeat, 4 hops

Error Probability \ Retransmission Procedure	Go-Back-N	Selective-Repeat
10^{-6}	89	>1000
10^{-5}	28	>1000
10^{-4}	8	236
10^{-3}	1	24

Table 3 Effect of the Number of Hops on the Performance of Error Recovery Schemes